



The Human Microbiome Project: A Community Resource for the Healthy Human Microbiome

Citation

Gevers, Dirk, Rob Knight, Joseph F. Petrosino, Katherine Huang, Amy L. McGuire, Bruce W. Birren, Karen E. Nelson, Owen White, Barbara A. Methé, and Curtis Huttenhower. 2012. The Human Microbiome Project: a community resource for the healthy human microbiome. PLoS Biology 10(8): e1001377.

Published Version

doi:10.1371/journal.pbio.1001377

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:10578922>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

The Human Microbiome Project: A Community Resource for the Healthy Human Microbiome

Dirk Gevers¹, Rob Knight^{2,3}, Joseph F. Petrosino^{4,5,6}, Katherine Huang¹, Amy L. McGuire⁷, Bruce W. Birren¹, Karen E. Nelson⁸, Owen White⁹, Barbara A. Methé^{8*}, Curtis Huttenhower^{1,10*}

1 The Broad Institute of MIT and Harvard, Cambridge, Massachusetts, United States of America, **2** Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado, United States of America, **3** Howard Hughes Medical Institute, Boulder, Colorado, United States of America, **4** Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas, United States of America, **5** Molecular Virology and Microbiology, Baylor College of Medicine, Houston, Texas, United States of America, **6** Alkek Center for Metagenomics and Microbiome Research, Baylor College of Medicine, Houston, Texas, United States of America, **7** Center for Medical Ethics and Health Policy, Baylor College of Medicine, Houston, Texas, United States of America, **8** J. Craig Venter Institute, Rockville, Maryland, United States of America, **9** Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, Maryland, United States of America, **10** Biostatistics, Harvard School of Public Health, Boston, Massachusetts, United States of America

The Human Microbiome Project (HMP) [1,2] is a concept that was long in the making. After the Human Genome Project, interest grew in sequencing the “other genome” of microbes carried in and on the human body [3,4]. Microbial ecologists, realizing that >99% of environmental microbes could not be easily cultured, developed approaches to study microorganisms in situ [5], primarily by sequencing the 16S ribosomal RNA gene (16S) as a phylogenetic and taxonomic marker to identify members of microbial communities [6]. The need to develop corresponding new methods for culture-independent studies [7,8] in turn precipitated a sea change in the study of microbes and human health, inspiring the new term “metagenomics” [9] both to describe a technological approach—sequencing and analysis of the genes from whole communities rather than from individual genomes—and to emphasize that microbes function within communities rather than as individual species. This shift from a focus on individual organisms to microbial interactions [10] culminated in a National Academy of Science report [11], which outlined challenges and promises for metagenomics as a way of understanding the foundational role of microbial communities both in the environment and in human health.

Pioneering medical microbiologists applied these approaches, finding far more microbial diversity than expected even in well-studied body site habitats [12]. Technological advances further

enabled sequencing of communities across the human body, and immunologists began exploring the fundamental role of microorganisms in the maturation of the innate and adaptive immune systems. Initial metagenomic studies of human-associated microbial communities were performed using the traditional Sanger platform [13,14]. Upon introduction of pyrosequencing [15], the number of 16S-based data sets increased dramatically [16,17]. The time was right to invest in a concerted study of the microbial communities associated with the human body and the metabolic capabilities they provide—the human microbiome (Figure 1) [18].

To coordinate these efforts relating the microbiome to human health, the NIH Common Fund launched the HMP as a community resource program (<http://commonfund.nih.gov/hmp/>) [19]. One of its main goals was to create a baseline view of the healthy human microbiome in five major areas (airways, skin, oral cavity, gastrointestinal tract, and vagina) and to make this resource available to the broad scientific community. Charac-

terizing the baseline state of the microbiota is a critical first step in determining how altered microbial states contribute to disease (e.g., [13,20–23]). Previous work showed wide inter- and intra-personal diversity of human-associated microbes [24], necessitating analysis of a large number of subjects and characterization of many reference bacterial genomes [25] to assist in interpretation of metagenomic data. The scope of the HMP thus required a particularly diverse consortium (Figure 2A), and collaboration among these teams ultimately stimulated research growth throughout the field and produced a study including the first consistent sampling of many clinically relevant body habitats, within a large population, with paired 16S profiling and deep metagenomic sequencing coverage for hundreds of microbial communities.

The HMP required careful consideration of ethical, legal, and social implications (ELSI) unique to the study of the microbiome [26]. Such research raises questions regarding traditional distinctions between self and non-self, human

Citation: Gevers D, Knight R, Petrosino JF, Huang K, McGuire AL, et al. (2012) The Human Microbiome Project: A Community Resource for the Healthy Human Microbiome. *PLoS Biol* 10(8): e1001377. doi:10.1371/journal.pbio.1001377

Published: August 14, 2012

Copyright: © 2012 Gevers et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This research was supported in part by National Institutes of Health grants U54HG004969 to B.W.B., D.G. and K.H.; U54HG004973 to J.F.P.; U54AI084844 to B.A.M. and K.E.N.; U01HG004866 to O.R.W.; R01HG005969 to C.H.; R01HG004872 to R.K.; Army Research Office grant W911NF-11-1-0473 and National Science Foundation grants NSF DBI-1053486 to C.H.; R.K. is an HHMI Early Career Scientist. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abbreviations: 16S, 16S ribosomal RNA gene; DACC, HMP Data Analysis Coordination Center; ELSI, Ethical Legal and Social Implications; HMP, Human Microbiome Project

* E-mail: chuttenh@hsph.harvard.edu (CH); bmethe@jvci.org (BAM)

The Community Page is a forum for organizations and societies to highlight their efforts to enhance the dissemination and value of scientific knowledge.

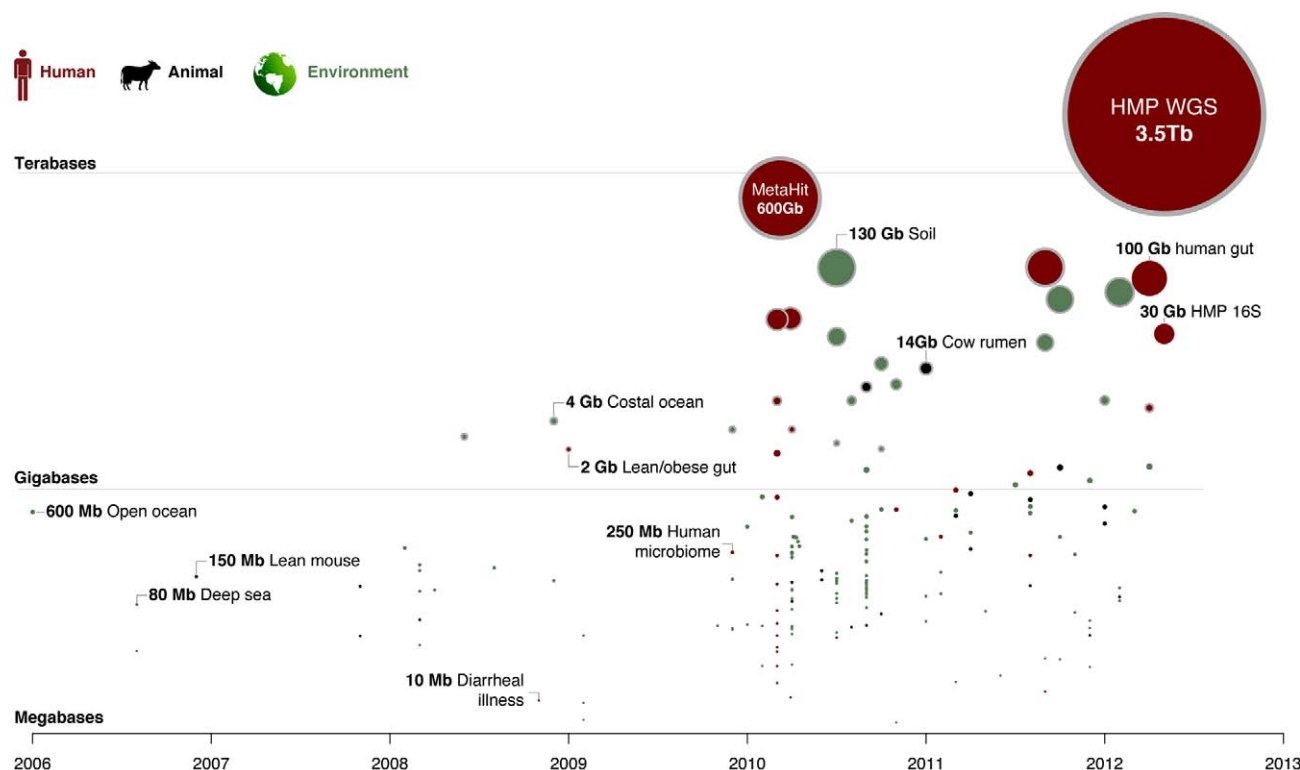


Figure 1. Timeline of microbial community studies using high-throughput sequencing. Each circle represents a high-throughput sequence-based 16S or shotgun metagenomic bioproject in NCBI (May 2012), indicating the amount of sequence data produced for each project (circle area and y-coordinate) at the time of publication/registration (x-coordinate). Projects are grouped by human-associated (red), other animal (black), or environmental (green) communities, and shotgun metagenomic projects are marked with a grey band. Selected representative projects are labeled: open ocean [68], deep sea [69], lean mouse [70], diarrheal illness [71], costal ocean [72], lean/obese gut [53], human microbiome [56], MetaHIT (gut) [58], cow rumen [73], soil (NCBI BioProject PRJNA50473), and human gut [74]. Note that HMP has deposited a total of 7.44 terabases of shotgun data in SRA, of which 49% is host DNA derived data that was filtered and only available through protected access in dbGaP project phs000228.

doi:10.1371/journal.pbio.1001377.g001

and non-human, genetics and environment, and health and disease. The prospect of manipulating the microbiota in ways that could permanently alter an individual's biological identity requires the development of new ethical paradigms analogous to, but not identical to, those already considered for gene therapy. Likewise, just as gene patents have proven controversial, defining who "owns" a microbiome raises difficult questions of intellectual property. The ELSI team helped to develop an appropriate sample collection protocol, to draft a template for informed consent, and consulted on ethical issues arising during the study, such as the possibility that unique human microbiome "signatures" [27] might compromise participant privacy. A portion of the HMP's dedicated research budget continues to be committed to integrating multidisciplinary approaches (including philosophical, social science, and legal methods) to study these issues and involve stakeholders including study participants, scientists, policy makers, patients, and indigenous populations.

Planning for Human Microbiome Studies: Tools, Techniques, and Design

Any study of human populations must put both subject protection and study design first, and the HMP was no exception. Power calculations for microbiome studies in human cohorts are particularly challenging, as they must simultaneously address assay types (e.g., 16S versus shotgun), depth of sequencing, taxon detection, and fold abundance changes in clades, genes, or pathways of interest [28–31]. After study design, as the HMP spanned multiple sequencing centers over a prolonged duration, the group established standardized and benchmarked protocols for sample collection [2], handling, and subsequent 16S profiling [32]. Metagenomic library construction was likewise standardized among centers, and stringent quality control was aided by the optimization of 16S read processing [33] and by improved taxonomic frameworks for classification of microbial sequences prior to biological interpretation [34].

Finally, quality data generation from appropriately designed microbiome studies enables a variety of subsequent computational analyses (Figure 2B). While we refer the reader to existing broader reviews of human microbiome bioinformatics [35–37], here we highlight numerous recent approaches specifically developed during the HMP. Several of these focused on microbial interactions, such as ecological network reconstruction [38,39]. Other computational methods dealt with metagenomic sequences, including both assembly-based [40,41] and assembly-free analyses of microbial community membership [42] and metabolic function [43]. Both data types enable taxonomic and phylogenetic profiling [44,45], and ecological metrics proved to associate microbial, gene, and pathway diversity on an unprecedented scale [2]. The HMP Data Analysis Coordination Center (DACC, <http://hmpdacc.org>) hosts all available HMP data and many tools, focusing the tremendous quantity of raw data through lenses such as SitePainter [46]; IMG/HMP, an HMP-specific version of the

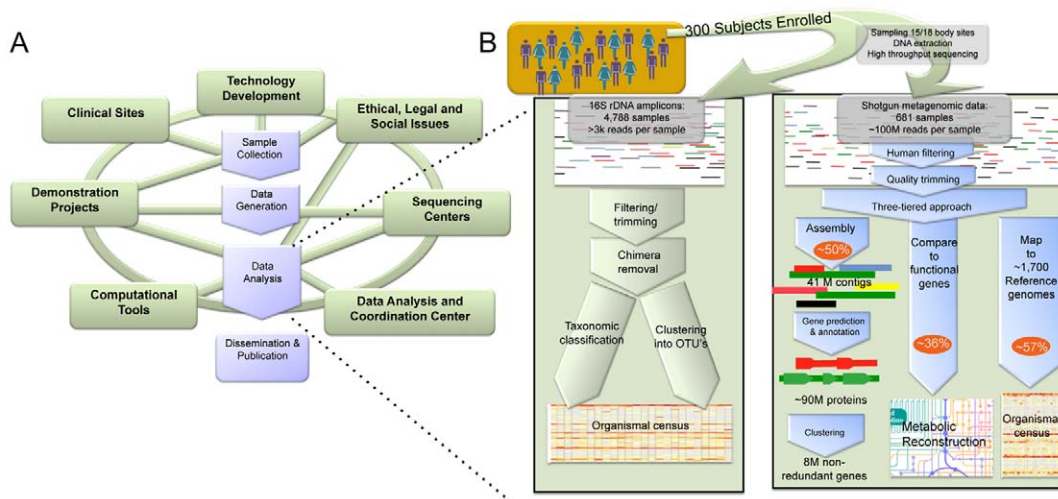


Figure 2. HMP consortium activities as a model for microbiome data generation and analyses. (A) Initiatives within the HMP coordinated to isolate samples, generate data, perform analysis, and publish results. *Technology development* was employed to develop novel bacterial culture and DNA isolation techniques. *Ethical Legal and Social Implications (ELSI)* work anticipated societal implications and guided policies associated with human subject microbiomes. *Clinical sites* were collected samples from large cohorts of healthy individuals, with nucleotide sequence information derived at four *sequencing centers* at the Baylor College of Medicine (BCM), the Broad Institute, the J. Craig Venter Institute (JCVI), and the Washington University Genome Institute (WUGI). Additional *demonstration projects* assessed primarily microbiome alterations related to disease. In addition to analysis throughout the HMP consortium, *computational tools* were funded to address, for example, genome assembly, microbial ecology, and statistical modeling. A *data analysis and coordination center* provided a portal to all data generated. (B) Overview of the analysis approaches that were the ultimate product of the HMP consortium, corresponding to data products and protocols available at <http://hmpdacc.org>. doi:10.1371/journal.pbio.1001377.g002

Integrated Microbial Genomes (IMG [47]) system; METAREP [48]; and MG-RAST [49], and efforts are ongoing to provide these data for meta-analysis alongside other human microbiome studies in the cloud.

Community Structure, Function, and a “Core” Human Microbiome

The HMP was designed in part to address a key question about our microbial selves: do all humans have an identifiable “core” microbiome of shared components comparable to our shared genome [50]? Several definitions of “core” have been proposed, recently unified in one conceptual framework [51]. Earlier studies reported that different people shared few microbes in their gut and skin microbiota [17,52–56], a greater fraction of their oral microbiota [56,57], or might be classifiable into multiple core microbiomes based on vaginal [20] and gut communities [58]. The HMP provides a comprehensive picture of the human microbiome covering multiple body sites and thus an in-depth exploration of these concepts. The study confirmed high inter-individual variation [59] and showed that even rare organisms in these communities are important reservoirs of genetic diversity [60].

Additionally, the large HMP cohort shows that the composition of the gut microbiome rarely clusters subjects into discrete types, as was suggested before on more limited data [61]; although other habitats such as the vagina can exhibit such clustering [20], the gut was most often characterized by smooth abundance gradients of key organisms [2].

A potentially more universal “core” human microbiome emerged during the consideration of microbial genes and pathways carried throughout communities’ metagenomes. While microbial organisms varied among subjects as described above, metabolic pathways necessary for human-associated microbial life were consistently present, forming a functional “core” to the microbiome at all body sites [2,43,53]. Although the pathways and processes of this core were consistent, the particular genes that implemented them again varied. Microbial sugar utilization, for example, was enriched for metabolism of simple sugars in the oral cavity, complex carbohydrates in the gut, and glycogen/peptidoglycan degradation in the vaginal microbiome [62]. The healthy microbiome may thus achieve a consistent balance of function and metabolism that is maintained in health, but with fine-grained details personalized by genetics, early life events, environmen-

tal factors such as diet, and a lifetime of pharmaceutical and immunological exposures [41].

The Healthy Microbiome Informs Studies of Disease

Data from individuals without overt signs of disease serve as an excellent reference for disease-associated microbiome studies, while also providing a comprehensive baseline for comparison of Western populations with disparate geographic, ethnic, and genetic cohorts [63]. The adoption of uniform sampling, nucleic acid extraction, sequencing, and analysis protocols is an important step in such integration, with some success already realized in, for example, several aspects of autoimmune disease. The inflammatory bowel diseases have long been linked to the human gut microbiome [22], with integration of host genotype, gene expression, and microbial membership now suggesting mediation of specific host-microbial interactions by human gene products as well as by host environment [64,65]. Bacteria are of course not the only mediators of dysbiotic disease, and metagenomic approaches can also be used to identify potential viral etiologies (e.g., in pediatric fever of undefined origin [66]). Likewise the “healthy” microbiome pro-

vides a baseline not only for integration with disease-related studies, but for broader populations such as a recent comparison using HMP protocols among a cohort of pregnant women [67]. The normal variation of the microbiome within healthy states and its potential misregulation in disease is thus being pursued in earnest, as related laboratory and computational methods continue to be adapted to better characterize the impact of bacteria, archaea, viruses, and fungi throughout human body habitats.

The HMP has thus greatly advanced our knowledge of the microbes in a healthy adult reference population, and provided much-needed infrastructure in

terms of reference genomes, laboratory protocols, computational methods, and ELSI considerations [1,2] to help enable a vast range of studies that will likely find associations between human-associated microbial communities and disease. The next steps will be to discover which of these microbial community changes result from disease and which cause it, to understand how healthy variation relates to variation within the context of different disorders, and to use a combination of laboratory and computational techniques to begin unraveling causal mechanisms on levels ranging from the molecular to the societal. In particular, the study of individuals of all ages and across cultures,

together with prospective longitudinal studies and careful work in *in vitro* and animal models, will be critical to developing both the science and the technology that will allow us to alter our microbial genomes, far easier to alter than the host genome within each of our “human” cells, in order to maintain and improve health.

Acknowledgments

The authors would like to thank Jean McEwen and Lita Proctor for consultation to confirm the accuracy of Human Microbiome Project program elements.

References

1. The Human Microbiome Project Consortium (2012) A framework for human microbiome research. *Nature* 486: 215–221.
2. The Human Microbiome Project Consortium (2012) Structure, function and diversity of the healthy human microbiome. *Nature* 486: 207–214.
3. Relman DA, Falkow S (2001) The meaning and impact of the human genome sequence for microbiology. *Trends Microbiol* 9: 206–208.
4. Davies J (2001) In a map for human life, count the microbes, too. *Science* 291: 2316.
5. Stahl DA, Lane DJ, Olsen GJ, Pace NR (1984) Analysis of hydrothermal vent-associated symbionts by ribosomal RNA sequences. *Science* 224: 409–411.
6. Pace NR (1997) A molecular view of microbial diversity and the biosphere. *Science* 276: 734–740.
7. Handelsman J (2004) Metagenomics: application of genomics to uncultured microorganisms. *Microbiol Mol Biol Rev* 68: 669–685.
8. Stein JL, Marsh TL, Wu KY, Shizuya H, DeLong EF (1996) Characterization of uncultivated prokaryotes: isolation and analysis of a 40-kilobase-pair genome fragment from a planktonic marine archaeon. *J Bacteriol* 178: 591–599.
9. Handelsman J, Rondon MR, Brady SF, Clardy J, Goodman RM (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem Biol* 5: R245–R249.
10. DeLong EF (2002) Microbial population genomics and ecology. *Curr Opin Microbiol* 5: 520–524.
11. National Research Council (US) Committee on Metagenomics (2007) The new science of metagenomics.
12. Kroes I, Lepp PW, Relman DA (1999) Bacterial diversity within the human subgingival crevice. *Proc Natl Acad Sci U S A* 96: 14547–14552.
13. Ley RE, Backhed F, Turnbaugh P, Lozupone CA, Knight RD, et al. (2005) Obesity alters gut microbial ecology. *Proc Natl Acad Sci U S A* 102: 11070–11075.
14. Gill SR, Pop M, Deboy RT, Eckburg PB, Turnbaugh PJ, et al. (2006) Metagenomic analysis of the human distal gut microbiome. *Science* 312: 1355–1359.
15. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, et al. (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437: 376–380.
16. McKenna P, Hoffmann C, Minkah N, Aye PP, Lackner A, et al. (2008) The macaque gut microbiome in health, lentiviral infection, and chronic enterocolitis. *PLoS Pathog* 4: e20. doi:10.1371/journal.ppat.0040020
17. Fierer N, Hamady M, Lauber CL, Knight R (2008) The influence of sex, handedness, and washing on the diversity of hand surface bacteria. *Proc Natl Acad Sci U S A* 105: 17994–17999.
18. Lederberg J (2000) Infectious history. *Science* 288: 287–293.
19. Peterson J, Garges S, Giovanni M, McInnes P, Wang L, et al. (2009) The NIH Human Microbiome Project. *Genome Res* 19: 2317–2323.
20. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, et al. (2011) Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci U S A* 108 Suppl 1: 4680–4687.
21. Baumgart M, Dogan B, Rishniw M, Weitzman G, Bosworth B, et al. (2007) Culture independent analysis of ileal mucosa reveals a selective increase in invasive *Escherichia coli* of novel phylogeny relative to depletion of Clostridiales in Crohn's disease involving the ileum. *ISME J* 1: 403–418.
22. Frank DN, St Amand AL, Feldman RA, Boedeker EC, Harpaz N, et al. (2007) Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc Natl Acad Sci U S A* 104: 13780–13785.
23. Young VB (2012) The intestinal microbiota in health and disease. *Curr Opin Gastroenterol* 28: 63–69.
24. Ursell LK, Clemente JC, Rideout JR, Gevers D, Caporaso JG, et al. (2012) The interpersonal and intrapersonal diversity of human-associated microbiota in key body sites. *J Allergy Clin Immunol* 129: 1204–1208.
25. Nelson KE, Weinstock GM, Highlander SK, Worley KC, Creasy HH, et al. (2010) A catalog of reference genomes from the human microbiome. *Science* 328: 994–999.
26. McGuire AL, Colgrove J, Whitney SN, Diaz CM, Bustillos D, et al. (2008) Ethical, legal, and social considerations in conducting the Human Microbiome Project. *Genome Res* 18: 1861–1864.
27. Fierer N, Lauber CL, Zhou N, McDonald D, Costello EK, et al. (2010) Forensic identification using skin bacterial communities. *Proc Natl Acad Sci U S A* 107: 6477–6481.
28. Knight R, Jansson J, Field D, Fierer N, Desai N, et al. (2012) Unlocking the potential of metagenomics through replicated experimental design. *Nat Biotechnol* 30: 513–520.
29. Kuczynski J, Lauber CL, Walters WA, Parfrey LW, Clemente JC, et al. (2012) Experimental and analytical tools for studying the human microbiome. *Nat Rev Genet* 13: 47–58.
30. Fitzner K, Heckinger E (2010) Sample size calculation and power analysis: a quick review. *Diabetes Educ* 36: 701–707.
31. Bacchetti P (2010) Current sample size conventions: flaws, harms, and alternatives. *BMC Medicine* 8: 17.
32. Jumpstart Consortium Human Microbiome Project Data Generation Working Group (2012) Evaluation of 16S rDNA-based community profiling for human microbiome research. *PLoS ONE* 7: e39315. doi:10.1371/journal.pone.0039315
33. Schloss PD, Gevers D, Westcott SL (2011) Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLoS ONE* 6: e27310. doi:10.1371/journal.pone.0027310
34. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, et al. (2012) An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J* 6: 610–618.
35. Kunin V, Copeland A, Lapidus A, Mavromatis K, Hugenholtz P (2008) A bioinformatician's guide to metagenomics. *Microbiol Mol Biol Rev* 72: 557–578, Table of Contents.
36. Hamady M, Knight R (2009) Microbial community profiling for human microbiome projects: tools, techniques, and challenges. *Genome Res* 19: 1141–1152.
37. Wooley JC, Godzik A, Friedberg I (2010) A primer on metagenomics. *PLoS Comput Biol* 6: e1000667. doi:10.1371/journal.pcbi.1000667
38. Friedman J, Alm E (2012) Inferring correlation networks from genomic survey data. *PLoS Comput Biol*. In press.
39. Faust K, Sathirapongsasuti JF, Izard J, Segata N, Gevers D, et al. (2012) Microbial co-occurrence relationships in the human microbiome. *PLoS Comput Biol* 8: e1002606. doi:10.1371/journal.pcbi.1002606
40. Koren S, Treangen TJ, Pop M (2011) Bambus 2: scaffolding metagenomes. *Bioinformatics* 27: 2964–2971.
41. Rho M, Wu YW, Tang H, Doak TG, Ye Y (2012) Diverse CRISPRs evolving in human microbiomes. *PLoS Genet* 8: e1002441. doi:10.1371/journal.pgen.1002441
42. Martin J, Sykes S, Young S, Kota K, Sanka R, et al. (2012) Optimizing read mapping to reference genomes to determine composition and species prevalence in microbial communities. *PLoS ONE* 7: e36427. doi:10.1371/journal.pone.0036427
43. Abubucker S, Segata N, Goll J, Schubert AM, Izard J, et al. (2012) Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Comput Biol* 8: e1002358. doi:10.1371/journal.pcbi.1002358
44. Segata N, Waldron L, Ballarín A, Narasimhan V, Jousset O, et al. (2012) Efficient metagenomic microbial community profiling using unique clade-specific marker genes. *Nat Methods*. Epub

ahead of print 10 June 2012. doi:10.1038/nmeth.2066

45. Liu B, Gibbons T, Ghodsi M, Treangen T, Pop M (2011) Accurate and fast estimation of taxonomic profiles from metagenomic shotgun sequences. *BMC Genomics* 12 Suppl 2: S4.
46. Gonzalez A, Stombaugh J, Lauber CL, Fierer N, Knight R (2012) SitePainter: a tool for exploring biogeographical patterns. *Bioinformatics* 28: 436–438.
47. Markowitz VM, Chen IM, Palaniappan K, Chu K, Szeto E, et al. (2012) IMG: the Integrated Microbial Genomes database and comparative analysis system. *Nucleic Acids Res* 40: D115–D122.
48. Goll J, Thiagarajan M, Abubucker S, Huttenhower C, Yooseph S, Methé BA (2012) A case study for large-scale human microbiome analysis using JCVI's metagenomics reports (METAREP). *PLoS ONE* 7: e29044. doi:10.1371/journal.pone.0029044
49. Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, et al. (2008) The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 9: 386.
50. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, et al. (2007) The human microbiome project. *Nature* 449: 804–810.
51. Shade A, Handelsman J (2012) Beyond the Venn diagram: the hunt for a core microbiome. *Environmental Microbiology* 14: 4–12.
52. Tap J, Mondot S, Levenez F, Pelletier E, Caron C, et al. (2009) Towards the human intestinal microbiota phylogenetic core. *Environmental Microbiology* 11: 2574–2584.
53. Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, et al. (2009) A core gut microbiome in obese and lean twins. *Nature* 457: 480–484.
54. Grice EA, Kong HH, Renaud G, Young AC, Bouffard GG, et al. (2008) A diversity profile of the human skin microbiota. *Genome Res* 18: 1043–1050.
55. Grice EA, Kong HH, Conlan S, Deming CB, Davis J, et al. (2009) Topographical and temporal diversity of the human skin microbiome. *Science* 324: 1190–1192.
56. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JL, et al. (2009) Bacterial community variation in human body habitats across space and time. *Science* 326: 1694–1697.
57. Zaura E, Keijser BJ, Huse SM, Crielaard W (2009) Defining the healthy “core microbiome” of oral microbial communities. *BMC Microbiology* 9: 259.
58. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, et al. (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464: 59–65.
59. Huse S, Ye Y, Zhou Y, Fodor AA (2012) A core human microbiome as viewed through 16S rRNA sequences clusters. *PLoS ONE* 7: e34242. doi:10.1371/journal.pone.0034242
60. Li K, Bihan M, Yooseph S, Methé BA (2012) Analyses of the microbial diversity across the human microbiome. *PLoS ONE* 7: e32118. doi:10.1371/journal.pone.0032118
61. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, et al. (2011) Enterotypes of the human gut microbiome. *Nature* 473: 174–180.
62. Cantarel BL, Lombard V, Henrissat B (2012) Complex carbohydrate utilization by the healthy human microbiome. *PLoS ONE* 7: e28742. doi:10.1371/journal.pone.0028742
63. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, et al. (2012) Human gut microbiome viewed across age and geography. *Nature* 486: 222–227.
64. Zhang T, Desimone RA, Jiao X, Rohlf FJ, Zhu W, et al. (2012) Host genes related to paneth cells and xenobiotic metabolism are associated with shifts in human ileum-associated microbial composition. *PLoS ONE* 7: e30044. doi:10.1371/journal.pone.0030044
65. Li E, Hamm CM, Gulati AS, Sartor RB, Chen H, et al. (2012) Inflammatory bowel diseases phenotype, *C. difficile* and NOD2 genotype are associated with shifts in human ileum associated microbial composition. *PLoS ONE* 7: e26284. doi:10.1371/journal.pone.0026284
66. Wylie KM, Truty RM, Sharpton TJ, Mihindukulasuriya KA, Zhou Y, et al. (2012) Novel bacterial taxa in the human microbiome. *PLoS ONE* 7: e35294. doi:10.1371/journal.pone.0035294
67. Aagaard K, Riehle K, Ma J, Segata N, Mistretta TA, et al. (2012) A metagenomic approach to characterization of the vaginal microbiome signature in pregnancy. *PLoS ONE* 7: e36466. doi:10.1371/journal.pone.0036466
68. DeLong EF, Preston CM, Mincer T, Rich V, Hallam SJ, et al. (2006) Community genomics among stratified microbial assemblages in the ocean's interior. *Science* 311: 496–503.
69. Sogin ML, Morrison HG, Huber JA, Mark Welch D, Huse SM, et al. (2006) Microbial diversity in the deep sea and the underexplored “rare biosphere”. *Proc Natl Acad Sci U S A* 103: 12115–12120.
70. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, et al. (2006) An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 444: 1027–1031.
71. Nakamura S, Maeda N, Miron IM, Yoh M, Izutsu K, et al. (2008) Metagenomic diagnosis of bacterial infections. *Emerg Infect Dis* 14: 1784–1786.
72. Gilbert JA, Field D, Swift P, Newbold L, Oliver A, et al. (2009) The seasonal structure of microbial communities in the Western English Channel. *Environmental Microbiology* 11: 3132–3139.
73. Hess M, Sczyrba A, Egan R, Kim TW, Chokhawala H, et al. (2011) Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. *Science* 331: 463–467.
74. Yilmaz P, Kottmann R, Field D, Knight R, Cole JR, et al. (2011) Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIXS) specifications. *Nat Biotechnol* 29: 415–420.